FILOLOGIJA 60, Zagreb 2013

UDK 811.16:811.17
Izvorni znanstveni članak
Primljen 5.XII.2013.
Prihvaćen za tisak 20.I.2014.

Ranko Matasović
Filozofski fakultet u Zagrebu
Odsjek za lingvistiku
I. Lučića 3, 10000 Zagreb
rmatasov@ffzg.hr

SUBSTRATUM WORDS IN BALTO-SLAVIC

This paper presents an analysis of those words, attested in Balto-Slavic, that do not have a clear Indo-European etymology and that could have been borrowed from some substratum language. It is shown that Balto-Slavic shares most of those words with other Indo-European languages of Northern and Western Europe (especially with Germanic), while lexical parallels in languages of Southern Europe (Greek and Albanian) are much less numerous. Georg Holzer's "Tememathic" hypothesis is also discussed, and a number of alternatives to his etymologies are suggested. It is argued that Balto-Slavic contains very few words borrowed from substratum languages that are not present in other branches of Indo-European.

Introduction

This could be a very short paper, indeed, since, as will become apparent, I will argue that there are no substratum words in Balto-Slavic. Having said that, I must make myself more precise: of course I don't mean that there are no substratum words in Baltic and Slavic languages. There are certainly plenty of them, although, as we will see, discovering them is a very difficult matter. What I will claim is that there is no layer of substratum words that can be posited for Proto-Balto-Slavic, and that is not shared by other branches of Indo-European, especially by Germanic, Celtic, and Italic.

Methodologically, it is very difficult to show that a set of loanwords from some unknown source *does not* exist. This would be tantamount to showing that a set of words in an Indo-European language, or group of languages, does not have a plausible Indo-European etymology, and that

075_102_Filologija_60_Matasovic_1.indd 75 19.01.2014 22:09:33

positing some unknown source is more probable. The problem lies in the definition of the concept "plausible Indo-European etymology". There are so many PIE roots, and their reconstructed meanings are often very, very general, so that it is nearly always possible to come up with a proposal to the effect that any word in any language is derived from a PIE root, and that there is a set of — often complicated — semantic changes by which the attested form is derivable from PIE. Take the following example.

PSl. **bára* 'shallow water expanse' (Croat. *bȁra,* Slov. dial. *barjè*, Bulg. *bára*, Russ. CSl. *bára*, Ukr. *bar* 'wet area between two hills', Pol. *barzyna,* dial. *barzówka*, Cz. dial. *bara*, *bařina,* Polabian *poro* 'mud', Vasmer I:53, Sławski I:191, ESSJa I:153, Matasović 2007). This word has been considered a borrowing from "Illyrian" (Gluhak 125), but this is improbable in the light of its attestations in West Slavic and in Ukrainian (not to mention the fact that we know next to nothing about "Illyrian"). It has also been connected to Gr. *bórboros* "mud, mire, filth" (Vasmer I:53), but this is formally difficult, since the Slavic forms are acute, as if from a laryngeal root. Moreover, Gr. *bórboros* can be plausibly connected to Arm. *kork* 'dirt' (Beekes 226f.). Rather, one is tempted to derive this Slavic word from PIE *b^heh_2- 'shine' (LIV. s. v.), cf. Ved. *bhā́ti* 'shines', Gr. *phaínō* 'shine'. The deverbal adjective from this root would have been *b^heh_2-ro-, and the collective > feminine noun *$b^heh_2reh_2$ 'shiny stuff', from which we can derive PSl. **bára* quite regularly. But how convincing is this etymology from the semantic point of view? There is no generally accepted method for judging semantic acceptability, and the only way to avoid complete subjectivity is to point to cases of parallel semantic development. This is indeed possible, in this case, as PIE *b^holHto- 'white' (Lith. *báltas*) yielded PSl. *bálto 'mud, swamp' (Croat. *blȁto,* Russ. *bolóto,* Cz. *bláto*, Pol. *błoto*). But, to be completely honest, I am not even convinced this etymology is as solid as usually assumed. So it is uncertain whether it can serve as a reliable parallel to the semantic development assumed for *bára.

Clearly, we need a method for establishing the likelihood that a given word is not inherited from a proto-language. I do not believe I can propose a generally acceptable method for this, but I believe that in discussing possible loanwords from unknown sources one has to bear in mind the following criteria (Matasović 2012):

a. Loanwords usually belong to semantic fields that are especially prone to borrowing (e.g. technological and cultural terminology, names of plants and animals)
b. The loanwords should be identifiable by their unusual phonolog-

075_102_Filologija_60_Matasovic_1.indd 76 19.01.2014 22:09:38

ical characteristics (i.e. unusual root structure, the presence of the rare PIE vowel *a, etc.); they should be isolated in the lexicon and not easily derivable from verbal roots.

c. We should avoid root etymologies. That is, reducing an attested form to a PIE root does not amount to a sound etymology, unless we can also explain its word-formation and relate it to established Indo-European patterns. If this cannot be done persuasively, the word is suspect of being a borrowing from some substratum language.

None of these criteria is sufficient by itself. Schrijver (1997) plausibly shows that in many cases we can posit a word in PIE even if some of the criteria mentioned above are violated. However, it is the "cumulative evidence rather than an individual criterium that tips the balance" (Schrijver 1997:296). Even so, claiming that a word, or a set of words in an IE language, are borrowings from some unknown substratum, often amounts to little more than saying that we do not know their etymology. Therefore Schrijver proposes an additional criterion, "namely the identification of substratum words by the fact that they show phonological and morphonological alternations which are regular in the sense that they recur in more than one etymon according to a certain pattern but irregular in the sense that they cannot be explained, for some reason or other, on the basis of Indo-European phonology and morphology" (Schrijver 1997:297). In what follows, we will see that there are many words of unclear, possibly substratum origin in Baltic and Slavic, that some of these words may indeed be attributable to Proto-Balto-Slavic, but that such words do not share any sort of regularity that could point to a common substratum as a source.

Indo-European Substratum in Balto-Slavic?

It is *a priori* not improbable that there were unknown Indo-European languages spoken between the Italic, Celtic and Germanic languages in the West, and Baltic and the Slavic languages in the East. These may have included "Pannonian" (Anreiter 2001), Venetic, Dacian, and several others about which we know next to nothing (Katičić 1976). It is not unlikely, moreover, that there are loanwords from such languages in Balto-Slavic, and the only question is whether they can be recognized as such. A reasonable attempt at discovering these substratum words was made by Georg Holzer (1989). He argued that there is a layer of Balto-Slavic vocabulary that had been borrowed from an unknown IE language, which

075_102_Filologija_60_Matasovic_1.indd 77 19.01.2014 22:09:38

he tentatively identifies with the language of the Cimmerians, a people of unknown origin who invaded Greece in the 7th century BC. According to Holzer, this language was characterized by a series of sound laws, different from Balto-Slavic sound laws, that can be observed in 45 different etymons attested in Balto-Slavic languages. The most prominent of these putative sound laws are the change of PIE mediae aspiratae into tenues, and the change of the PIE tenues into mediae (hence the handy name of the substratum language, *Temematic*).

A number of Holzer's "Tememathic" etymologies appears very plausible at first sight. Thus, Holzer derives PSl. *svobodъ 'free', *svoboda 'freedom' from PIE *swe-poti, *swe-poteh$_2$, with the first element *swe- from the root of the reflexive pronoun (cf., e.g., OCS *svojь* 'own'), and the second element from PIE *poti- 'master, lord' (Skr. *páti-*, Gr. *pósis* 'husband', etc.). However, the Slavic words have a perfectly viable etymology that does not resort to "Tememathic" sound laws: we can derive *svobodъ from PIE *swo-b^{h}o- (cf. OCS *svobьstvo* 'person', OPr. *subs* '-self, own', Asg. *subban*, Goth. *sibja* 'kin', Latv. *at-svabinât* 'set free') with the rare, but well attested, suffix *-oda (cf. e.g. OCS *agoda* 'fruit' vs. Lith. úoga 'id.', Vasmer III:596, Trautmann 291, or *lьgoda 'lightness' > Croat. *lagoda*, cf. *lьgъkъ 'light' > Russ. *légkij*).

Likewise, Holzer derives PSl. *těsto 'dough' from PIE *d^{h}oyg'^{h}-to- by "Tememathic" sound laws (cf. Eng. *dough,* Germ. *Teig*, Gr. *teíkhos* 'wall', etc.), but there is no need for this, since the Slavic word is plausibly connected with Gr. *staĩs* 'flour of spelt mixed and made into dough', OIr. *táis* 'dough', W *toes* 'dough' < PIE *teh$_2$is-to-, cf. also OHG *theismo* 'dough' (EDPC 374, Derksen 492f.).

In spite of the fact that his "Tememathic" hypothesis has not met with much critical reception, not to speak of general acceptance,[1] it is our opinion that it cannot be dismissed out of hand. Holzer has offered a viable solution to a number of Slavic (to a lesser extent also Baltic) etymological problems, although it is difficult to accept his overall hypothesis of a single substratum language to which all of the etymologically difficult Balto-Slavic words discussed in his book should be ascribed. For our purposes it is important to note that only 11 out of the 45 tememathic roots he reconstructs have cognates in both Baltic and Slavic; this is under 25 % of the total number. Another problem is that in all but one case (Tem. *k'ey-

[1] Generally speaking, Holzer's hypothesis was better accepted among Slavic scholars (e.g. Moszyński 1992, Brozović 1992) than among Indo-Europeanists (however, Kortlandt, who belongs both categories of scholars, reviewed Holzer's work rather positively, see Kortlandt 2003).

075_102_Filologija_60_Matasovic_1.indd 78 19.01.2014 22:09:38

ro- 'orphan') the Baltic and Slavic reflexes of the Tememtic roots do not go back to a single prototype. This means that most of his etymologies are root etymologies, which should not be admitted easily.

Let us review Holzer's "Tememtic" etymologies of words that are attested in both Baltic and Slavic:

Tm. *berg'-, *borg'-: Lith. *bìržė, biržė, biržìs,* Latv. *bìrze* 'furrow', OCS *brazda*, Russ. *borozdá,* Croat. *brázda,* Cz. *brázda*, etc. from PIE *pork'o- (Lat. *porca*, Gall. *rica*, OHG *furuh*). The Baltic and Slavic forms cannot be derived from the same prototype, so the derivation from Tememtic amounts to a root etymology. An alternative etymology (Vasmer I:109) relates the Slavic words to PIE *b^hers- 'point' (OHG *burst* 'bristle', OIr. *barr* 'top'); PIE *b^hors-d^heh_2 would regularly yield PSl. *borzda, but in this case the Baltic words must be unrelated. Smoczyński (61—2) derives them from the PIE word for 'birch' (*b^herHg'^{h}- > OCS *brěza*, Germ. *Birke*, etc.). It would originally have denoted a furrow delineated by birch branches, which is semantically difficult.

Tm. *delno- / *dolno- / *dolgo-: *dolnъ 'palm of the hand' (OCS *dlana,* Russ. dial. *dolón',* Croat. *dlȁn*, Pol. *dłoń*), Lith. *délna, délnas*, Latv. *dęl̃na*. This set of words is derived by Holzer from PIE *telH- (Skr. *tála-* 'plain, surface', OCS *tъlo* 'ground, floor'). However, a better etymology connects these words to PIE *del(h_1)- 'hew, cut' (e.g. Smoczyński 99, cf. also IEW 194, Lith. *délti, deliù,* Lat. *dolo* 'cut into shape', *dolabra* 'knife', OIr. *delb* 'form', EDPC 95, Alb. *dalloj* 'cut', perhaps also Lith. *dalýti* 'divide'). The original meaning would have been 'carved, rasped surface'. In Russian dialects, reflexes of *dolnъ mean also 'threshing-floor' (ESSJa V:63—64[2]). Note, however, that OIr. points to a root without the laryngeal, which is difficult to square with the accentuation of the BSl. forms, which were acuted.

Tm. *g'ebi / *g'obi: PSl. *zobъ 'oats' (Croat. *zôb,* Russ. *zob*, Pol. *zób*), Lith. žebìkė, žėbìkė 'sack for cattle food', Latv. *zebenieks* 'id.' These words are derived by Holzer from the same root as German *Hafer,* OHG *habaro* 'oats', apparently from PIE *k'op- by "Tememtic" sound laws. However, the Germanic words for 'oats' can be plausibly connected with OIr. *corca*, W *ceirch* 'oats' and derived from a proto-form *kork^{w}ro- (Kluge 347), and the Balto-Slavic words can be related to PSl. *zobati 'peck', Russ. dial. *zobat'*, OPol. *zobać*, Croat. *zòbati* 'peck, eat grains', Lith. žẽbti 'eat dry substances, gobble' (Derksen 547).

[2] ESSJa's etymology connecting these BSl. words with Gr. *thénar* 'palm of the hand', OHG *tennar* 'id.' is far-fetched and formally very difficult.

075_102_Filologija_60_Matasovic_1.indd 79 19.01.2014 22:09:38

Tm. *k'ey-ro-: PSl. *sirъ 'orphaned' (Russ. *síryj*, Cz. *sirý*), Lith. šeirỹs 'widower', cf. Av. *saē-* 'orphaned', which points to PIE *k'ey-. Smoczyński (628) and EWA II:615 adduces also Ved. śayú- 'orphan' < *k'eyu-. Holzer (134f.) relates these words to Skr. *hīyate* 'is left', *hāni-* 'lack', Gr. *khēros* 'orphaned, empty', *khḗra* 'widow', Lat. *hērēs* 'heir'. However, the PIE root can be safely reconstructed as *g'^{h}eh$_1$-, and Skr. *hī-* contains the reflex of the laryngeal in the interconsonantal position. The verbal root is Skr. *hā-* 'leave'. There is no evidence for PIE *i- in this root, so the etymology proposed by Holzer cannot be correct.

Tm. *ponto-, *pontā / *ponti-: PSl. *pąto 'fetter' (OCS pl. *pąta,* Croat. *pȕto*, Russ. *púto*, Pol. *pęto*, Derksen 417), Lith. *pántis* 'rope (for fettering horses)', *pánčia* 'id.', OPr. *panto* 'fetter'; Holzer derives this from PIE *b^{h}end^{h}- 'bind' (= Eng. *bind*, Skr. *bandh-*, etc.), but an alternative etymology is readily available: these BSl. words come from PIE *(s)penH- 'stretch, spin, weave' (Arm. *henowm* 'weave', Lith. *pinù* 'attach', OHG *spannan,* Gr. *pénomai* 'get tired', IEW 988). The same root is attested in PSl. *pęta* 'heel', OPr. *pentis* 'id.', Lith. *péntis* 'back side of an axe' (Smoczyński 450).

Tm. *proko- / *pirkā: PSl. *prokъ 'remaining' (ORuss. *prokъ* 'rest', OCS *prokъ*, OPol. *prokny* 'every', Derksen 421), Lith. *pirkià, pirkčià* 'peasant house, cottage' , dial. *pirkáitė* 'larder, pantry'; these words are derived by Holzer from PIE *b^{h}rg^{h}o-'keep, preserve' (Russ. *beregú*). Again the Slavic and Baltic words do not match exactly, so we are dealing with a root etymology. A simpler solution within Slavic is to derive *prokъ from the preposition *pro- 'forward' and a suffix *-kъ which is common in adjectives (Snoj 582). Cf. also Gr. *próka* 'immediately' and Lat. *procul* 'far'.

Tm. *swep-/ *sup-: PSl. *svepetъ 'moving around', OCS *svepiti sę* 'move' Cz. *svapato* 'bee-hive' (Derksen 475), Lith. *supù, sùpti* 'move with difficulty, to rock, cradle', PSl. *sъtъ 'bee-hive'; these are derived by Holzer from PIE *web^{h}- 'weave' (OHG *weban*, Gr. *hyphaínō*, etc.). However, a rather plausible etymology derives these words from PIE *sewp- 'to strew, to throw', cf. also Lat. *supo* 'pour, strew', ORuss. *sъpъ* 'embankment, hill', OCS *suti* 'pour, strew'. This hypothesis, admittedly, involves *Schwebeablaut.* PSl. *sъtъ can be derived from *sup-to- (Vasmer II:702 thinks this etymology is uncertain).

Tm. *tel- 'calf' (PSl. *telę 'calf' > Croat. *tèle,* Russ. *telënok,* Pol. *cielę*), Latv. *tęleñs,* dial. *teļš,* Lith. dial. *tẽlis, tẽlias*); derived by Holzer from PIE *d^{h}eh$_1$-l- 'sucking' (Gr. *thẽlys* 'feminine'). However, Snoj derives this from a root noun *tēl, *tlés from the root meaning 'to bear' (Latv. *iz-tilt* 'bring', Lat. *tollo, tollere* 'pick up', Gr. *tlẽnai* 'bear, support', IEW 1060f.), but this is se-

075_102_Filologija_60_Matasovic_1.indd 80 19.01.2014 22:09:38

mantically difficult. A connection is possible with *toliti 'calm, soothe', but again the semantic connection is weak. Most likely, this Slavic word is of Turkic origin, cf. Tuvan *tel* 'calf', Kazakh *tel*, Yakut *tīl*. Starostin (*www.starling.ru*) reconstructs Proto-Turkic *Tē.l- 'a kid or calf sucking two milch-ewes or cows'. Since there are no direct Turkic loanwords in Baltic, it is probable that the Baltic words for 'calf' were borrowed from Slavic.

Tm. *trono-, *tronto- 'drone' (Croat. *trût,* Russ. *trúten',* Pol. obs. *trąt,* Derksen 498), Lith. *trãnas,* Latv. *trans* vs. Germ. *Drohne* and Gr. *thrõnaks*). These words indeed appear to be of substratum origin, since the initial *t- in BSl. cannot correspond regularly to Germanic *d- and Greek *th-. Note, however, that the Slavic and Baltic formations are not identical, so they may have been borrowed from different sources, or through different intermediaries.

Tm. *twer- / *tur- / *tworo-: Lith. *tveriù*, *tvérti* 'take, hold', Latv. *tvêru, tvêrt*, OPr. *turei* 'has', OCS *za-tvoriti* 'close', Lith. *tvártas* 'stable'. Holzer derives these words from PIE *d^hwer- 'close' (OCS *dvorъ* 'courtyard', *dvьrь* 'door'), but a perfectly viable PIE etymology is available, cf. Gr. *sóros* 'urn', *seiré* 'rope', 'cord (with a noose)' (Smoczyński 698). LIV accepts this and reconstructs the root as *twerH- 'seize'. Cf. also Lith. *tvorà* 'fence', OCS *tvarь* 'creature', 'creation' with the lengthened grade (PIE *twōrH-, which may have been a root noun).

Tm. *twirdo- / *twirto- 'strong', 'firm', 'solid': PSl. *tvьrdъ, Lith. *tvìrtas,* Latv. *tvîrts*; Holzer derives these words from PIE *d^hwer-to- 'having doors', but they are quite obviously related to the preceding etymon. The semantic development was from 'seized' to 'squeezed' and 'firm, solid'. In any case, the Baltic and Slavic suffixes do not match.

Everything considered, Holzer's hypothesis remains unproven, especially if it is meant to show that his "Tememe" loanwords were borrowed during the Balto-Slavic period. It remains possible that both Slavic and Baltic borrowed independently from some unknown Indo-European language, which may or may not be identical with Holzer's "Tememe".

Other possible loanwords from Indo-European substratums in Balto-Slavic were sought in the words showing *Palatalwechsel*, where Baltic and/or Slavic show plain velars, while evidence from other Satem-languages shows that a palatalized velar must be posited for PIE (Gołąb 1972, 1990). However, the examination of the evidence shows that in most cases regular depalatalizations in Balto-Slavic can be posited, e.g. after *s-mobile, or before a resonant followed by a back vowel (Matasović 2005). In the few

075_102_Filologija_60_Matasovic_1.indd 81 19.01.2014 22:09:38

cases where phonological depalatalizations cannot be assumed,[3] it is always simpler to assume that a word was borrowed from a known group of languages (especially Germanic and Celtic) than from some unknown substratum.

Non Indo-European Vocabulary in Balto-Slavic

Of all the language families in the world about which we have any knowledge, Uralic is the only one that is a reasonable candidate for the donor of substratum vocabulary in Balto-Slavic. Today, nearly all Uralic languages are spoken near the Slavic and/or Baltic speaking area, and Uralic, Baltic, and Slavic contacts certainly stretch back deep into prehistory. However, although there are many Uralic loanwords in individual Balto-Slavic languages, especially in Russian and Latvian, there do not seem to be any Uralic loanwords that could be attributed to Proto-Slavic, or to Balto-Slavic periods (Kallio 2005).

Of course, there may have been other, now extinct non-IE languages and/or language families in Europe, and Balto-Slavic may have borrowed words from any of them. It has long been known that Indo-European languages spoken in Europe share many vocabulary items that do not have cognates in the Asian branches of Indo-European. Some of this vocabulary may have been preserved from Common PIE only in the west, but some is likely to have been borrowed from unknown substratum languages of Europe. These languages may have been genetically related, if they were all descended from the language of the first Neolithic farmers who had arrived in Europe from Asia Minor in the 7th and 6th millennium BC, but at least equally possible is that they belonged to different language families, some of which were perhaps related to Basque, while others were not. These languages, or language families, may ultimately have been spoken by the Mesolithic hunter-gatherers who adopted agriculture from the settlers arriving in Europe from Asia Minor and the Middle East. The next sections contain a list of words in Baltic and Slavic languages that have cognates only in European branches of IE.[4]

[3] E.g. in the Slavic word for "goose" (PSl. *gąsь > Russ. *gus'*, Pol. *gęś* vs. Lith. žąsìs, OPr. *sansy*, Latv. *zùoss*), which might be from Germanic (cf. OHG *gans*), and in the Slavic word for "cow" (PSl. *korva > Russ. *koróva*, Croat. *krȁva*), cf. W *carw* 'deer'.

[4] We will exclude words that have very limited distribution in Slavic, especially those that are limited to Russian, Baltic, and North Germanic, as there is no evidence that such words go back to Proto-Slavic. A good example is Russ. *sig* 'a kind of salmon, *Coregonus lavaretus*', which was presumably borrowed from the same unknown source as Lith. *sykis*, Latv. *sīka, sīga*, and OIc. *sīkr*, Swedish *sik* (Orel 330, Vasmer II:621).

075_102_Filologija_60_Matasovic_1.indd 82 19.01.2014 22:09:39

The northwestern connections

In this section we will look at those words that Balto-Slavic shares with Northwestern Indo-European dialects (especially Germanic, Celtic and Italic) that are suspect of having substratum origin.[5]

PSl. *olьxa 'alder' (Russ. *ol'xa,* Pol. *olcha,* Bulg. *elxá*, ESSJa VI:23—25.), Lith. *aĨksnis, eĨksnis*, dial. *alìksnis* 'alder' (Smoczyński 11); there are parallels in Germanic (OE *alor,* OHG *elira* < *alisa, Orel 15) and Latin (*alnus* 'alder', de Vaan 34—35). The variation in the Anlaut points to probable substratum origin (thus Derksen 370—1).

PSl. *bagno 'swamp' (Russ. dial. *bagnó,* Pol. *bagno*, Cz. *bahno*, ESSJa I:125—127), Dutch *bagger* 'mud', OHG *bah* 'stream', OIc. *bekkr* 'brook, rivulet', OE *bece* (< PGerm. *bakiz, Orel 33), perhaps MIr. *búal* 'flowing water' (if from *boglo-). Long *a in Slavic could be due to Winter's law, in which case this is a very early loan from some unknown source.

PSl. *balъka 'ravine, pool' (Pol. dial. *bałka* 'pool, pond', Russ. dial. *balka* 'ravine', perhaps also Croat. dial. *balka, bala* '*Stipa pennata L*', ESSJa I:149), Lith. *balà* 'swamp', Latv. *bala* 'woodless valley', OE *pōl* 'pool', OHG *pfuol* 'id.' < PGerm. *pōlaz (Orel 292). The derivation from PIE *b^helH- 'white' or *$b^h eh_2$- 'light' (ESSJa) is impossible because of initial *p- in Germanic. If these words are related, it is probably a substratum word[6].

PSl. *belenъ, *bьlnъ 'henbane' (Russ. dial. *belen,* OPol. *bleń,* Cz. *blín,* Bulg. *blen* 'day-dream', ESSJa I:185—187); these words are certainly related to the Germanic words for 'henbane', e.g. OE *beolone,* OS *bilene,* perhaps Gaul. *belenountían* (Asg.). In Germanic we also find forms with different suffixes that probably point to substratum origin (Schrijver 1999), cf. OHG *bilisa* (Germ. *Bilsenkraut*), Swedish *Bolmört*, Kluge 111).

PSl. *bobъ 'bean' (Russ. *bob*, Croat. *bòb,* Pol. *bób,* ESSJa II:148), OPr. *babo* 'bean' (perhaps a Slavic loanword), OE *bean*, OHG *bōna* (< PGerm. *baunō < *bab-nō), Lat. *faba* 'bean'. The usual comparison with Gr. *phakós* 'lentil' and Alb. *bathë* '*Vicia faba*' does not lead anywhere (Demiraj 94, see also Kuiper 1995).

PSl. *drągъ 'bar, pole' (OCS *drągy* 'sticks', Pol. *drąg*, Croat. *drûg* 'rail'),

[5] There are, of course, many other lexemes that Baltic, Slavic (or both) share with Northwestern IE dialects, especially with Germanic (see Stang 1972), but there are no reasons to believe they are of substratum origin (according to the criteria sketched in the Introduction).

[6] Smoczyński (42) separates Lith. *balà* from PSl. *balъka and relates it to the Lith. adjective *bãlas* 'pale', 'white', which is presumably related to *báltas* 'white' (in turn from the same root as PSl. *bolto 'mud', see above).

075_102_Filologija_60_Matasovic_1.indd 83 19.01.2014 22:09:39

Lith. *drañgas* 'pole (used as a lever)', ESSJa V:129—130; OIc. *drangr* 'detached pillar of rock', OIc. *drengr* 'heavy stick'; a variant with *-k- is PSl. *drąkъ 'bar, pole' (Russ. *druk*, Cz. *drouk*, Croat. *druk*), ESSJa V:130—131. The alternation of voiced and voiceless root-final stops might point to the substratum origin (thus Derksen 121).

PSl. *goląbь 'pigeon' (Croat. *gȍlūb,* OCS *goląbъ*, Russ. *gólub'*, Pol. *gołąb,* ESSJa VI:215—17), is too similar to Lat. *columba* for the similarity to be accidental; the forms may go back to *Kol-omb^{h}-, with the suffix *-omb^{h}-/*-emb^{h}- which is common in bird-names (see also Derksen 2000). Holzer (1989) considers the Slavic form a "Temematic" loanword. Baltic does not show direct reflexes, but cf. Lith. *gelumbė̃* 'blue cloth', OPr. *golimban* 'blue' vs. Russ. *golubój* 'pale blue', OPr. *gołęby* 'greyish, blue-grey' (ESSJa VI:217).

PSl. *grabrъ 'hornbeam' (Croat. *grȁb,* Serb. dial. *gàbar, gràbar*, Sln. *gâber, grâber*, Russ. *grab*, Pol. *grab*, Cz. *habr,* ESSJa VII:99—100). Cf. OPr. *wosigrabis* '*Evonymus Europaeus*'. The appurtenance of Umbr. *Grabovius* (epithet of Jove on *Iguvine Tablets*) and Ancient Macedonian *grãbin* 'a kind of tree' is uncertain. The comparison with Lith. *skrõblas* 'hornbeam' (Smoczyński 568), perhaps Lat. *carpinus* 'id.'[7] shows that we may be dealing with a substratum word, reconstructable as *(s)grāp/b^{h}-, a distinctively non-IE shape.

PSl. *grъměždь, *krъmeljь 'fester in the corners of the eyes, gramiae' (Russ. CS *grъměždь,* Croat. *krȍmēlj*, Slov. *krmẹ́lj*, ESSJa VII:158—9). The similarity with lat. *grāmiae, gramiae* 'viscous humour, rheum in the corner of the eyes' and Goth. *qrammiƕa* 'moisture' is too great to ignore, but these words cannot be reduced to a single prototype, so they were probably borrowed from some non-IE source (thus also De Vaan 270).

OLith. *kasulas* 'hunting spear', Lat. *corulus* 'hazel-tree' < *kosulo-, OIr. *coll*, W *coll* < PCelt. *koslo- 'hazeltree', OHG *hasal*, OIc. *hasl* < PGerm. *xaslaz 'id.' (Orel 164). The alternation in the shape of the suffix (*-slo- / *-sulo-) points, perhaps, to non-IE origin.

PSl. *elbedь, *olbądь 'swan' (Russ. *lébed'*, Cz. *labut'*, Croat. *lȁbūd,* OPol. *łabęć,* Pol. *łabędź*, ESSja VI:19, XXXII:50—51.), OHG *albiz, elbiz*, Lat. *olor*, OIr. *elu*, W *alarch,* pl. *eleirch*. The alternation of *e- and *o- in Slavic is an instance of Rozwadowski's change, but still the vocalism of the different words for 'swan' is difficult to square with a PIE prototype (De Vaan 427). Another trace of substratum origin could be the alternation of stem-final *-d and *-t. Celtic and Italic point to *el-, Germanic to *al- (*h$_2$el-), and Slavic can be from both, but the acute on the root implies that the root end-

[7] De Vaan 94 relates this to Hitt. *karpina-* 'kind of fruit tree', Lith. *skir̃pstas* 'elm' and speculates that the root is the same as in Lat. *carpo* 'pluck, gather' (PIE *ker-p-).

075_102_Filologija_60_Matasovic_1.indd 84 19.01.2014 22:09:39

ed in a laryngeal. Perhaps the forms with *a- were influenced by *h_2elb^ho- 'white' (> Lat. *albus*).

PSl. *xvoja 'needles of a coniferous tree', (Pol. *choja*, Russ. *xvója,* Pol. *choja,* ESSJa VIII:125—126), Lith. *skujà*, Latv. *skuja* (Smoczyński 568, doubting the etymology), OIr. *scé, scíad* [Gsg.] 'thorn bush' (PCelt. *skwiyat-, EDPC 339). We may posit an original alternation *skwoy- / *skuy-, but the shape of the root is distinctively non-Indo-European.

PSl. *klęčati 'kneel' (OCS *klęčęšta* [Ndu part. pres. act.] 'kneeling, Pol. *klęczeć*, Croat. *kléčati,* ESSJa X:28—9), Lith. *klénkiu* 'walk with difficulty'; Snoj (277) derives these words from PIE *kleng-, *klenk- (Lat. *clingere*, MHG *lenken,* MHG *gelenke*, OHG *(h)lanca* 'thigh', 'haunch', Kluge 310). Note, however, that the reflexes are attested only in the European branches, where the alternation of *k and *g may point to substratum origin.

PSl. *klenъ 'maple' (Russ. *klën*, Pol. *klon,* Croat. *klèn,* Vasmer I:567), Lith. *klẽvas* (Smoczyński 297), Latv. *kļavas,* OE *hlīn*, OIc. *hlynr* < PGerm. *hluni. W *kelyn* 'holly', OCo. *kelin,* OIr. *cuilenn* (EDPC 213) come from PCelt. *kolino- and need not be related (likewise OE *holegn* 'holly', OHG *hulisboum*). The connection with Macedonian *klinó(s)trokhos* (Theophrastus) is also doubtful.

PSl. *kobьcь, *kobuzь, *kobъzь 'kite, hawk' (Russ. *kóbec,* , Pol. *kobiec,* Croat. *kòbac,* Russ. dial. *kóbuz,* Pol. *kobuz, kobz,* Vasmer I:582—3), OHG *habuh* 'kite', OE *hafoc* 'hawk', OIc. *haukr* < PGerm. *habukaz (Orel 138). The Slavic noun *kobьcь seems to be formed with the diminutive suffix *-ьcь, and the form *kobuzь may be directly comparable to the Germanic words for hawk. However, a reconstruction *kobug'o- does not look like a PIE word.

*PSl. *lęntьja 'lentil' (CSl. *lęšta*, Russ. *ljača*, Bulg. *léšta*, Croat. *léća,* Vasmer II:84, ESSJa XV:63—65), OHG *linsa* 'lentil' (Kluge 521 claims that this is not a Latin loanword), Lat. *lēns* 'lentil' (De Vaan 238). Gr. *láthyros* 'pulse, *Vicia sativa*' is probably unrelated.[8] Since lentils were originally cultivated in the Middle East and the Balkans, it seems probable that the word was borrowed from some Mediterranean source, perhaps one of the lost non-IE languages of the Balkans.

Lith. *mãkas* 'purse', PCelt. *makinā (MW *megin* 'bellows', Co. *mygen*, EDPC 254), OHG *mago* 'belly', OE *maga* 'id.' < PGerm. *magōn, Orel 253.

[8] Lith. *lęšis* 'lentil' is probably borrowed from Germ. *Linse* (Smoczyński 348).

075_102_Filologija_60_Matasovic_1.indd 85 19.01.2014 22:09:39

The vowel *a points to a non-IE origin.

Lith. *nõtrynė* 'nettle', OPr. *noatis,* perhaps also PSl. *natъ 'leafy top of a root vegetable' (Ukr. dial. nat', Pol. *nać*, Croat. *nât* ESSJa XXIII:186f.), PGerm. *natōn, *natilōn (OE *netele*, OHG *nazza, nezzila*, Orel 281f.), PCelt. *ninati- (OIr. *nenaid,* MW *dynat* [pl], EDPC 290). The reflexes point to *ninati- and *nāti-, a very un-Indo-European form.

PSl. *ovьsъ 'oats' (Russ. *ovës,* Pol. *owies*, Croat. *òvas,* Derksen 384f.), Lith. *avižà*, Latv. *àuza*, OPr. *wyse,* Lat. *avēna* < *aweksnā; both the alternation of the voiced and voiceless palatalized velars in BSl. and the unusual shape of the suffix (*-eKs- / *-iKs-) point to the substratum origin (Oettinger 2003. 189).

PSl. *rogъ 'horn' (OCS *rogъ,* Croat. *rȍg*, Russ. *rog*, Pol. *róg,* Derksen 438), Lith. *rãgas*, Latv. *ràgs,* OPr. *ragis.* Smoczyński 495 relates these words to Lith. *regėti* 'see', assuming a root *reg^{h}- 'to be visible, appear', Germ. *sich regen*. Snoj 627 also starts with the meaning 'to stick out' and connects Lith. *rogsóti* 'stick out', cf. also CS *rogozъ* 'bull-rush, sedge', Russ. *rogóz.* LIV reconstructs the root as *reg^{h}-, but there seems to be also a variant *rek- in PGerm. *rahō 'stick, pole' (ON *rá* 'sailyard pole', OHG *raha* 'shuttle', Orel 293), Lith. *rėklės* [pl.] 'scaffolding'. If these words are related (which is by no means certain), they show the alterantion of voiced and voiceless stops at the end of the first syllable, which might be indicative of substratum origin.

PSl. *ruxo 'clothes' (Russ. *rúxo*, Croat. *rȕho*, Pol. *rucho*) can be derived from *ruk-s-o- and related to PGerm. *rukkaz (OHG *rocko* 'overcoat', OE *rocc* 'upper garment', Orel 308), PCelt. ?rowkk- / *ruk- (MIr. *rucht* 'tunic', W *rhuch* 'film', 'layer', 'jerkin', 'coat', EDPC 315). The alternation between the geminate and the simple *-k- (and *x in Slavic) looks non-IE.

Lith. *serbentà* 'blackberry, redcurrant', dial. *sarbentà* (Smoczyński 543), Russ. dial. *sorbalína* 'blackberry, *Rubus fruticosus*', *serbalína, serbarína* 'rosehip', Byelorussian *cerbalín* 'blackberry' (Vasmer II:697), Lat. *sorbus* 'servicetree' (De Vaan 576), perhaps also Swedish *sarv* 'rudd, redeye, *Leuciscus erythrophthalmus*' (a reddish fish). Since the Slavic reflexes are limited to Russian dialects (and an isolated Byelorussian form), this might be a Baltic loanword in Slavic. The verbal root seems to be attested in Lith. *sir̃bti* 'ripen' (perhaps originally 'redden', since all of the berries denoted by this set of verbs are red).[9]

PSl. *sьrebro *'silver'* (OCS *sьrebro*, Russ. *serebró*, Pol. *srebro*), Lith. *sidã-*

[9] One is tempted to connect these words to OIr. *suib* 'strawberrry', W *syfi* < PCelt. *subi- (EDPC 358), which are without etymology.

075_102_Filologija_60_Matasovic_1.indd 86 19.01.2014 22:09:39

bras, Latv. *sidrabs* (Smoczyński 546, who claims that the Baltic forms were borrowed from Slavic), Goth. *silubr*. The appurtenance of Celtib. śilapur is probable, but not quite certain (EDPC 41).

OPr. *spurglis* 'sparrow', OE *spearwa* 'sparrow', PCelt. *sfrawo- 'crow' (Bret. *fraw,* Co. *frau*, EDPC 334), perhaps Lat. *parra* 'a kind of bird', Gr. (Hesych) *sparásion* 'sparrow-like bird', cf. Schrijver 1997:304.

*(i)vьlga 'oriole' (CS *vlъga*, Russ. ívolga, Pol. *wilga*, dial. *wywiołga, wiwielga*, Croat. *vùga*, Bulg. *avlíga,* ESSJa VIII:251—252), Lith. *volungẽ* 'oriole', Latv. *vãluôdze*, MHG *witewal* 'oriole', Dutch *wielewaal*. Since *wite-* of MHG *witewal* means 'wood', it has been claimed that the apparent prefix *i- in Slavic is from *iwo-, i.e. that *ivьlga is from *ivo-vьlga (Derksen 217), where *ivo- is the word for 'willow' (Russ. íva, Croat. ìva, Lith. íeva 'bird-cherry'). However, even if this is accepted, one cannot reconstruct a Balto-Slavic proto-form for the word for 'oriole', so borrowing from a substratum language and different adaptations in Baltic and Slavic seem very likely.

This set of words contains many nouns referring to cultural items, flora and fauna, that are readily borrowed in situations of intensive language contact. Otherwise, there are very few, if any, formal features that can be gathered from this material. One thing worth mentioning is that the alternation of the suffixes *-is- and *-(e)n- can be observed in a number of items referring to plant and tree names, e.g. PSl. *bel(e)nъ 'henbane' and OE *beolone* vs. OHG *bilisa*, PSl. *klenъ 'maple' vs. OHG *hulisboum*, and perhaps Lith. *alk̃snis* 'alder' and Lat. *alnus* vs. PSl. *olьxa, OHG *elira*.

The southern connections

The lexical items ascribable to a substratum and shared by Slavic and/or Baltic with the "southern" European languages (especially Greek and Albanian) are much less numerous, and they are generally based on less reliable etymologies:

PSl. *kolyba ,*koliba 'hut' (Slov. *kolíba*, Cz. dial. *kolíba* 'tent', Bulg. *kolíba*), Gr. *kalýbē*; possibly related to PSl. *xalupa 'hut, cottage' (Croat. *halupa* (Kastav), Slov. *halúpa*, Russ. dial. *xalúpa,* Pol. *chałupa*). PSl. *kolyba* may have been borrowed from Greek at a relatively late stage (after the Slavic migrations), but this does not solve the problem of the ultimate origin of this word.

PSl. *kosъ 'blackbird' (Croat. *kôs*, Russ. *kos*) is usually related to Gr. *kópsikhos, kóssyphos* (also *kóttyphos, kóssykos*) 'id.'; this is a very uncertain

075_102_Filologija_60_Matasovic_1.indd 87 19.01.2014 22:09:39

etymology, since the original form of the Greek word for 'blackbird' is difficult to ascertain. PSl. *kosъ would be derivable from *kopso-, but it could just as well be from PIE *kosó-, from the root *kes- 'to scratch, to comb' (OCS česati, etc.). The semantic development would have been from 'the scratcher' to 'blackbird', and it is interesting to note that blackbirds do indeed scratch the soil and litter to pull earthworms (admittedly, so do other species of birds). A semantic parallel development can be observed in Croat. čèšljugar ('goldfinch', *Carduelis caruelis*), which is derived from češalj 'comb', from the same PIE root *kes- (Skok I:311f.).

PSl. *kъlpь 'swan' Pol. dial. *kiełb* 'swan', Russ. dial. *kolp'* 'spoonbill', Croat. dial. *kûp* 'swan', ULus. *kołṕ.* These words are certainly related to Latv. *gùlbis*, Lith. *gulbė̃*, *gulbìs* (4), with initial *g-*, Derksen 261. The original form had *g-* to judge by OIr. *gulban* 'beak, sting', W *gylfin* 'beak, snout', if this etymology is correct (EDPC 168—169 does not accept it).

PSl. *mъrky 'carrot' (Russ. *morkóv*, Croat. *mrkva*, Pol. *marchew*), OHG *morha;* often related to Gr. *brákana* [pl.] 'wild herbs' (Vasmer II:158—159). The connection with the Greek word is very dubious, as the meanings do not match, and there is no evidence that Greek *br-* is from *mr-; Beekes (235) considers it Pre-Greek and doubts the connection with the putative Slavic and Germanic cognates. It is also possible that the Slavic word was borrowed from Germanic, but this does not solve the question of its ultimate origin.

PSl. *trъstь 'reed, cane' (OCS *trъstъ*, Russ. *trost'*, Pol. *treść*, Croat. *trst*), Lith. *trùšiai*, *trušìs*, Latv. *trusis*, Trautmann 330, Vasmer III:141. The connection with Gr. *thrýon* 'reed, rush' (IEW 1097, DELG 443) is possible only if one assumes the development *truso- > *truho- > *thruso- in Greek, which is dubious. The unusual Anlaut in Greek could also be used as an argument for borrowing from some non-IE language.

Pan-European substratum words

In a few cases, the substratum words found in Slavic have cognates both in the South (usually in Greek) and in the North (usually in Germanic); good examples are:

PSl. *bьrglězь 'finch' (Russ. *bergléz,* Croat. *brglijez*, Vasmer I:75), Gr. *frygílos*, Lat. *fringilla* (Vasmer I:75). However, Beekes (1593) notes that the meaning of the Greek bird name is not completely certain ('chaffinch' is only one possibility). It is also possible that all of these words are independent, originally onomatopoetic formations.

075_102_Filologija_60_Matasovic_1.indd 88 19.01.2014 22:09:39

PSl. *čemerъ ‘*Veratrum album*, hellebore’ (Russ. dial. čémer, Pol. dial. *czemier*, Croat. čȅmēr ‘venom, anger’, ESSJa IV:52—53), Lith. *kẽmeras*, OHG *hemera* ‘id.’, Gr. *kámaros*.

PSl. *čeremuxa, *čermъxa, *sermъša ‘ramson’, ‘bird cherry’ (Croat. *srȉjemuša, crȉjemša*, Russ. dial. čerëmuxa, Pol. *trzemcha*, Slov. črêmha, ESS-Ja IV:66—68), Lith. dial. *kermušẽ* ‘tip of a drill, ramson’, Gr. *krómyon, krémyon* ‘onion’, OIr. *crem* ‘wild garlic’, ‘leek’, W *craf* [Collective] ‘garlic’ < PCelt. *kremu-, *kramu- (EDPC 222), OE *hramsa* ‘ramson’, ESSJa IV:66—68. The forms with initial *s- < *k’ in Slavic show that this is a very early loan, since it exhibits *Palatalwechsel*. The comparison with Greek and Celtic forms does not allow the reconstruction of a PIE prototype, so this word was probably borrowed from some non-IE source.

PSl. *česnъ ‘garlic’ (Russ. česnók, Croat. čèsan, Pol. *czosnek*, ESSJa IV:89—90), PCelt. *kasninā ‘garlic’ (MIr. *cainnenn*, OW *cennin* [p] ‘leeks, daffodils’, OCo. *kenin* gl. allium, EDPC 193); the alternation *e/*a seems to point to substratum origin (Schrijver 1995:495), but the e-vocalism of Slavic may be due to the influence of the verbal root *kes- ‘to comb’, (?) ‘to peel’ (OCS česati).

PSl. *konopь ‘rope’ (Slov. *konȍp*, Bulg. *konóp*, Russ. dial. *konóp, konóp’* OPol. *konoṕ*, ULus. *konop*, Vasmer I:615), OHG *hanaf* ‘hemp’, OE *haenep* < PGerm. *hanapa- (Kluge 354), Gr. *kánnabis*. This is an old Wanderwort, perhaps of oriental origin, cf. Sumerian *kunibu* ‘hemp’.[10]

PSl. *kъlkъ ‘hip’ (Russ. dial. *kolk*, Croat. *kȕk*, Bulg. *kъlka*, Vasmer I:600), Lith. *kulkšnìs*, ‘ankle’, OPr. *culczi* ‘hip’, Latv. *kulksnis* ‘tarsal joint’; the comparison with Lat. *calx* ‘heel’, ‘hoof’ (De Vaan 86) seems very probable, but a common prototype cannot be reconstructed. The distinctly non-IE form of the root *kVlk- points to non-IE origin.

PSl. *makъ ‘poppy’ (Russ. *mak*, Croat. *màk*, Pol. *mak*, ESSJa XVII:149—151), Lith. dial. *māguonė*, Latv. *maguône*, OPr. *moke*. The Lithuanian and Latvian words were probably borrowed from Germanic, (Derksen 299—300), cf. OHG *mago* besides *māho*, and the OPr. word may be a loanword from Polish. The vowel alternation in Germanic is unclear, as the OHG forms point to *makō̆n- and *mā̆kōn.[11] The comparison with Gr. *mḗkon*, Dor. *mā̃kon* ‘poppy’ shows that the proto-form of the root *meh$_2$ko-, pehaps alternating with *mh$_2$k- is original. The vowel alternation in Ger-

[10] Herodotus (4.74—75) claims that hemp was imported from Scythia. Lat. *cannabis* is a Greek loanword.

[11] Boutkan (1998) notes that a similar alternation exists in PGerm. *magaÞ- ‘girl, maid’ and *mēg- > Goth. *megs* ‘son-in-law’, which he also considers as loanwords from some non-IE source.

075_102_Filologija_60_Matasovic_1.indd 89 19.01.2014 22:09:39

manic might be due to different adaptations of the same foreign sound (?/æ/) either as *a, or as *ē.

PSl. *rěpa 'turnip' (Russ. *répa*, Croat. *rèpa*, Cz. řepa), Lith. *rópė*, OHG *ruoba, ruoppa,* Lat. *rāpum,* Gr. *rhápys, rháphys*. The diverging vocalism of the words for 'turnip' in European languages suggests this is a loanword from some non-IE source (thus also De Vaan 14).

PSl. *roda 'heron' (OCS *roda*, Croat. *róda,* Skok III:163), Gr. *erōdiós, rhōdiós,* Lat. *ardea,* perhaps also OIc. árta 'a kind of duck'. The diverging vocalism of the words for 'heron' cannot be reconciled with a PIE reconstruction. This word was probably as migratory as the bird it denotes.

PSl. *versъ, *verskъ 'heather' (Russ. *véresk,* Cz. *vřes,* Croat. *vrìjes*, Derksen 516), Lith. *vìržis,* Latv. *vîrzis*, *vìrsis*; the Slavic form points to *werk'-, and the Baltic forms to *wrk'- or *wrg'-. T. Pronk (p. c.) thinks that Baltic -ž- may be due to influence from Lith. *veržti* 'string', 'tighten', 'squeeze' < *uerǵh-. Undoubtedly related are Gr. *ereíkē* (< *wer-eyk'-o-), OIr. *froech,* W *grug* (< *wroyk'o-), but no common prototype can be posited (Snoj assumes *werk'-, *wrg'-, *wereyk'-, *wroyk'-). Everything considered, this is probably a loanword from some Non-IE language (thus also Beekes 452).

Again, it should be noted that Baltic reflexes of words belonging to this stratum of the lexicon are often missing. Words belonging to this layer appear to be "Wanderworter", and their original source(s) and paths of transmission are probably impossible to determine.

Words limited to Balto-Slavic

Finally, there is a group of words that is attested only in Balto-Slavic, and there are no plausible cognates elsewhere in Indo-European.[12] Here is a tentative list:

BSl. *bauKura- > PSl. *bugorъ, *bugъrъ 'hillock' (Russ. and Ukr. *bugór*), Latv. *baũgurs* 'id.', cf. perhaps also Lith. *kaũb(u)ras* 'hillock' (with metathesis and unexplained *k-*); a connection with the PIE root *b^hewg^h- 'to bow, twist' (Goth. *biugan,* Skr. *bhuj-,* ESSJa III:79) is formally difficult (because of the acute in Latvian) and semantically not particularly attractive. Since no reflexes are attested in W and S Slavic, the E Slavic words may have been borrowed from Baltic.

[12] The source for this section is Trautmann's dictionary (1921); of course, it contains other lexemes that are limited (at least in terms of word-formation) to Balto-Slavic, but we have collected those that are most likely to be loanwords from some substratum language according to our criteria sketched in the Introduction to this paper.

075_102_Filologija_60_Matasovic_1.indd 90 19.01.2014 22:09:40

BSl. *Purna- > PSl. *bъrna 'snout' (Slov. dial. *bŕna* 'carnival mask depicting an animal', Croat. *brna*, Maced. *brna* 'nose-ring of animals', Lith. *burnà* (3) 'mouth, face', but Latv. *puŕns* 'snout'. The etymological connection with Arm. *beran* is conjectural (Derksen 69), as well as with the root *b^{h}erH- 'to bore' (Lat. *forāmen* 'opening'), not to speak of the connection with MIr. *bern* [ā f] 'gap, breach, pass'.

BSl. *darga- > PSl. *dorgъ 'dear' (Croat. *drâg,* Russ. *dorogój*, Pol. *drogy*), Latv. *dā`rgs*; Snoj (120) and SP IV, 121 tentatively relate this to the root *d^{h}reg^{h}- 'hold' (IEW 252, Av. *dražaite* 'holds', Gr. *drássomai* 'hold', OCS *drъžati*). The original meaning would have been 'powerful, (able to) hold'. ESSJa V:77 claims that Latv. *dārgs* is a Slavic loanword.

BSl.*e/a-mela- > PSl. *emela,*jьmela 'mistletoe' (Russ. *oméla*, Pol. *jemioła*, Croat. ìmela, dial. òmela), Lith. *ãmalas, ẽmalas,* Latv. *amuols, amuls*, OPr. *emelno*; cf. also (with Ablaut) Latv. *āmuls, āmulis* 'id.' Smoczyński (13) mentions the traditional etymology which relates these words with the root *h$_1$em- 'hold, take' (Lat. *emo*, OCS *imati*, Lith. *imù, imtĩ*), because mistletoe is used in the production of glue. More likely, only Slavic forms with initial *jь- (e.g. Croat. *imela*) were influenced by this root. Smoczyński also mentions the alleged connection with Lat. *amārus* 'bitter', but does not really believe it. ESSJa V:26 compares Eng. *mistel,* OHG *mistil* but offers no solution. Kluge (535f.) derives the Germanic forms from *mihs-tlo- and relates them to Gr. *iksós,* Lat. *viscum,* perhaps also PSl. *višnja 'cherry', but the initial *m- is unexplained. Could this be an instance of the prefixation of *a- found in Northwest European loanwords?

BSl. *graSa- 'threatening' > PSl. *groziti 'threaten' (Russ. *grozít', grožú,* Pol. *grozić*, Croat. *gròziti*, ESSJa VII:143), *groza 'horror' (OCS *groza,* Russ. *grozá,* Pol. *groza,* Croat. *gròza,* ESSJa VII:141—142), Lith. *gražóti* 'threaten', *gražùs* 'beautiful', OLatv. *grę̃zns* 'beautiful, luxurious'. With *-s-* we also have Lith. *grasìnti* 'threaten', *grasùs* 'threatening'. The connection with OIr. *gargg* 'wild', Gr. *Gorgṓ* 'Gorgon', *gorgós* 'terrible' (Snoj 193, IEW 353) is very dubious. The alternation of Slavic *s and Lith ž might point to a substratum origin, but ESSJa VII:141f. claims that the Lithuanian forms with -ž- were borrowed from Slavic and that the forms with *-s- are unrelated, which is probably true. Smoczyński (195f.) leaves Lith. *gražùs* without etymology.

BSl. *Kl(a)uša- > PSl. *gluxъ 'deaf' (OCS *gluxъ*, Russ. *glúxyj*, Croat. *glûh,* Cz. *hluchý*, ESSJa VI:146—147), Lith. *glušas* 'dumb'; Snoj (176) relates this to PSl. *glupъ 'silly' and *glumъ, *gluma 'joke'. Vasmer (I:277) compares Lith. *glusnùs* 'obedient', *klusnùs* 'id.', dial. *klùsas* 'dumbish'. The alterna-

075_102_Filologija_60_Matasovic_1.indd 91 19.01.2014 22:09:40

tion of *k- and *g- may be a sign that these words are of substratum origin. ESSJa's (VI:147) derivation from PIE *k'lows- 'hear' by "expressive" change of *k to *g does not lead anywhere.

BSl. *Krauša- > PSl. *grúša 'pear' (Russ. *grúša*, Cz. *hruška*, Pol. *grusza*), PSl. *kruša (Bulg. *krúša,* Croat. *krȕška*, dial. *krȕšva*, Pol. dial. *krusza*, ULus. *krušva*, LLus. *kruša*; in Baltic only with *k-, cf. Lith. *kriáušė*, OPr. [pl.] *crausios*, Vasmer I:314. ESSJa (VII:156) connects these words with the verbal root found in PSl. *grušiti / *krušiti 'crush', but this is hard to believe on semantic grounds.[13]

BSl. *īni- > PSl. *inьjь 'hoar-frost' (OCS *inii,* Russ. ínej, Cz. *jíní,* Croat. înje, ESSJa VIII. 235—6), Lith. ýnis. These words might perhaps be related to Germ. *Eis* < Germ. *īsaz (Orel 204), which may be, in turn, from the same root as Av. *aēxa-* 'cold', *isu-* 'icy'. The problem is that we would expect the cluster *-sn- to be preserved in Balto-Slavic. Snoj (202) considers the possibility that the original form was *ivьnьje and compares Slov. îvje 'hoar-frost', Russ. dial. íven' 'id.' ESSJa VIII:235 agrees with this and claims that Lith ýnis was borrowed from Slavic. According to ESSJa, *jьvьnьje would be derivable from *jьva, *jьvica 'edge', 'crust' (Serb. *ivica,* Bulg. *iva* 'edge of a cloth'), but this is semantically doubtful. Moreover, the etymology of *jьva, *jьvica is just as unknown as the etymology of *inьjь.

BSl. *ledu- > PSl. *ledъ 'ice', Lith. *lẽdas,* Latv. *lędus,* OPr. *ladis*; ESSJa 14:91—2 leaves this word without an etymology. Vasmer (25) adduces MIr. *ladg* 'snow' (MoIr. *laogh*), but this comparison is very uncertain. The MIr. word is poorly attested and both stem formation and inflection are uncertain (G sg. *ladga* or *laide* according to DIL).

BSl. *me/arGa- > PSl. *merža 'net', Lith. *márška* 'sheet', 'table-cloth, 'drag-net', Latv. *marȇga* 'railing', 'gallery'; *męrȇga* 'railing', 'gallery', *mārsna*; Smoczyński (374) doubts the connection because Lith. -šk- cannot correspond to Slav. -ž-. Snoj (420) reconstructs PIE *merHg^{h}- from the root *(s)mer- 'weave' (Gr. *mérmīs* 'thread', Hitt. *išmeri-*) and compares also Gr. *brókhos* 'sling', 'rope', 'noose', MIr. *braige* 'prisoner' (quoted from IEW 733); actually the word is *brága* 'captive', 'prisoner', 'hostage', which DIL treats as the development of OIr. *brágae* [f t] 'neck', 'throat'. Beekes finds evidence that both Gr. *mérmīs* (932) and *brókhos* (243) are Pre-Greek. Pronk (p. c.) derives the BSl. words for 'net' from PIE *merg'- 'border', 'limit'

[13] ESSJa points to a putative semantic parallel in Lat. *pirum* 'pear', which it derives from *peys- 'to beat, to crush' (OCS *pьxati,* etc.), but this is improbable, as Gr. ápion 'pear' shows that these words are borrowed from some non-IE substratum (with the mysterious prefix *a-* posited by Schrijver 1997?). Original PIE *piso- would be reflected as *peru- in Latin (De Vaan:467), cf. *sero* 'sow' < *sish$_1$oh$_2$.

075_102_Filologija_60_Matasovic_1.indd 92 19.01.2014 22:09:40

(OIr. *mruig*, Lat. *margō*, Goth. *marka* 'area, border', MoPers. *marz* 'region'. If the original meaning of the root is preserved in Hitt. *mārk-^{i}/mark-* 'divide', 'separate', the meaning of 'net' might be derivable from it. The acute in BSl. would be the result of Winter's law. However, the Balto-Slavic forms point to the root-final plain velar, while the words meaning 'border, region' show reflexes of a palatalized PIE guttural.

BSl. *pausta- > PSl. *pustъ 'empty, deserted' (OCS *pustъ*, Russ. *pustój*, Pol. *pusty*, Croat. *pûst*, Derksen 424), OPr. *pausto*; perhaps from *powH-d^{h}-to- 'cleansed' (LIV *pewH- (1), cf. Ved. *punãti*, OHG *fouwen* 'sieve')?

BSl. *(a)rayHša- > PSl. *orěxъ 'nut, walnut' (Russ. *oréx*, Pol. *orzech*, Croat. òrah, Derksen 374), Lith. *ríešutas*, Latv. *riẽksts*, OPr. *buccareisis* 'beech-nuts'; Snoj 476f. connects also Alb. *arrë*, Gr. pl. árya. Smoczyński (515f.) compares Gr. *ereíkō* 'break, tear, rend', but Beekes (452f.) relates this to Lat. *rixa* 'quarell', Lith. *riẽkti* 'cut hay' < PIE *h$_1$reyk-. The alternation of initial *o- (< *a-) in Slavic and the vowelless form in Baltic is similar to the alternation posited by Schrijver (1997) in words from NW European substratum.

BSl. *seyHlā > PSl. *sila 'force' (OCS *sila*, Russ. *síla*, Pol. *siła*, Croat. *sȉla*, Derksen 451), Lith. *síela* 'soul', OPr. *seilin* 'diligence'; perhaps to *seh$_2$i- 'bind', Lith. dial. *sienù*, Latv. *sìet*, Skr. *sā-*, Hitt. *išhiya*, cf. PCelt. *soyto- 'magic' (MW *hud*, EDPC 352) and OIc. *seiðr* 'magic, charm' from the same root. Snoj (655) compares OIr. *sethar* 'powerful', W *hydr* 'strong', 'bold', but this must be from a different root without the laryngeal (*seyd^{h}-), and OIr. *sethar* probably does not exist (it is a ghost-word). The connection with the root *seh$_2$i- is possible only under the assumption that the Slavic and Baltic forms are not etymologically identical, as Slavic *sila must be from *sih$_2$leh$_2$, while Lith. *síela* must reflect *seh$_2$ileh$_2$. This is quite improbable, so it is better to consider BSl. *seylā a loanword from some unknown source.

BSl. *śama- > PSl. *somъ 'sheat-fish' (Russ. *som*, Pol. *sum*, Croat. *sòm*, Derksen 461), Lith. šãmas, Latv. *sams*; the usual connection with Gr. *kamasẽnes* 'a kind of fish, [pl.]' is doubtful. Snoj (682) thinks of a connection with Croat. *smûđ* '*Lucioperca sandra*', Slov. *smûč*, as well as Russ. *sudák* 'id.', but this is difficult. Smoczyński 624 connects these words to Gr. *kámaks* 'shaft, pole', but the semantic connection is to weak (? 'fish as long as a pole').

BSl. *tranTa- > PSl. *trądъ 'tree fungus, tinder' (OCS *trądъ* 'illness', Pol. *trąd* 'leprosy', Croat. *trûd*, Lith. *trandìs* 'woodworm, moth', Latv. *trûdi* [Npl.] 'mould' (Derksen 497—8). A variant with final voiceless stop seems to be attested in Russ. *trut* 'tinder' and Bulg. *trъt*. Probably related to Lith. *trẽndu* 'be eaten by worms' (Vasmer III:144—145). The derivation from the root *terd- 'drill' (Skr. *tard-*, thus in LIV) is possible only if one assumes

075_102_Filologija_60_Matasovic_1.indd 93 19.01.2014 22:09:40

Schwebeablaut, or that the BSl. root *trend- was abstracted from the stem of the nasal present *tr-n-d-, for which there is no evidence.

BSl. *traupa- > *trupъ 'dead body', 'log', OPr. *trupis* 'log'; Snoj (789) compares Lith. *trupė́ti* 'crumble' and Gr. *trūpáō* 'bore, pierce trough'. Vasmer (III:143—144) adduces Latv. *trupêt* 'rot, decay', Lith. *trupùs* 'crushed to pieces', Latv. *trupe* 'soil, earth', Smoczyński 692—693 accepts this and adduces forms with initial *d-* such as Latv. *drupt* 'crumble', Latv. dial. *draûpît* 'crumble'. Beekes (1513) doubts the appurtenance of Gr. *trȳpáō* and Gr. *trýpē* 'hole' because these point to a laryngeal root, and there is no evidence for a laryngeal in BSl. Everything considered, it is safest to separate the Balto-Slavic words from the putative Greek cognates, and the alternation between initial *t- and *d- in Baltic might point to substratum origin.

BSl. *tulśi- > PSl. *tъlstъ 'thick' (Russ. *tólstyj*, Pol. *tłusty*, Croat. *tȕst*), Lith. *tuĨžti*, *tulštù* 'swell', perhaps Latv. *tulzis* 'gall', Lith. *tulžìs* (4)[14], Trautmann 332, Vasmer III:117. The alternation between *st* and ž might point to a substratum origin. There may be a connection to Germ. *tulguz 'firm, steadfast' (Goth. *tulgus*, Orel 411) and *talgō 'tallow' (Germ. *Talg*, OIc. *talg*, Orel 400), if we start from a substratum *TolK'- / *TlK'-.

Balto-Slavic also has a number of verbal roots which do not appear to have any cognates elsewhere. Apart from this fact, there is no reason to consider them borrowings from some unknown substratum. Here is a tentative list:

BSl. *k'auH- > PSl. *sovati 'shove' (Russ. *sovát'*, OCz. *suvati*, Slov. *suváti*, Derksen 462), Lith. šáuti 'shoot', Latv. šaũt, dial. *saũt*; Snoj (711) compares Skr. *suváti* 'pushes, sets in motion', Hitt. *šuwāi-* 'push away' (Kloekhorst 797f.), but this does not explain Lith. š-; Smoczyński 626—627 does not connect the BSl. words with PIE *sewH- but rather with *k'ewH- 'throw', which LIV reconstructs as BSl. only.

BSl. *pel-/ *pāl- 'burn' > PSl. *paliti 'burn' (OCS *paliti,* Croat. *páliti*, Pol. *palić,* Russ. *palít',* Derksen 390), *pepelъ 'ashes' (Croat. *pèpeo,* Russ. *pépel,* Pol. *popiół*), Lith. *pelenaĩ* 'ashes', OPr. *pelenne*; LIV 805. Related is also *pelinъ (Snoj 504). The root does not seem to be attested outside Balto-Slavic, but we think it might be derived by metathesis from PIE *leh$_2$p- (> *lāp- > BSl. *pāl-), cf. Hitt. *lāpzi* 'glows', Gr. *lámpō*, OPr. *lopis* 'flame', OIr. *lasaid* 'burn' (EDPC 235). In that case the words for 'ashes' must be from a different root, perhaps PIE *pelH- 'chaff' (OPr. *pelwo,* OCS [pl.] *plěvy*, etc., IEW 802).

[14] Smoczyński 694 derives these words for 'gall bladder' from *źulti- by metathesis.

075_102_Filologija_60_Matasovic_1.indd 94 19.01.2014 22:09:40

BSl. *sewt- 'be mad' > PSl. *šutъ 'fool' (Russ. šut), *šustrъ 'movable, handy' (Russ. šústryj, Vasmer III:439—440), Lith. *siùsti, siuntù* 'be mad', Latv. šust, šutu 'be angry', Vasmer III:439—440, Trautmann 269. Perhaps related are the reflexes of PSl. *žustrъ 'quick', 'handy' (Croat. žȕstar, Byelorussian žostry 'handy', Gluhak 714).

Looking at the set of words adduced in the previous section, we must now ask if they exhibit any sort of quasi-regular, but non-IE morphological or phonological patterns, which would point to their substratum origin. Do they show any of the patterns typical of the North West European substratum established by the Leiden school Indo-Europeanists? These include the alternation of voiced (and aspirated?) and voiceless stops at the end of the first syllable (Boutkan 2003), especially between *p, *b, *bb, *ff, *pp and *mp, in Germanic (Kuiper 1995); the prefixing of a word-initial *a- and the reduction in the vocalism of the remainder of the word in Italic, Celtic, and Germanic (Schrijver 1997), the presence of the non-IE vowel *-a- (Kuiper 1995, Beekes 1996, Boutkan 2003) and its alternation with *-ai- in Germanic (Schrijver 1997), as well as the clusters *kl- and *kn- in Germanic. Boutkan (2003) also mentions the typically disyllabic root-shape (CVCVC) with alternating vowels in the second syllable. None of these features are found in our material, except for the alternation of voiced and voiceless stops and fricatives (and, in the case of the word for 'nut', perhaps the alternation of initial *a- with zero).[15]

Within Baltic, words with this alternation are quite common. A list can be found in Endzelīn 1971:75—76. Here is a selection:

Lith. *viskėti* / *vizgėti* 'swing',
Lith. *virpėti* / *virbėti* 'vibrate'
Latv. *klēpis* 'armful', Lith. *klėbỹs*
Latv. *drupas* 'ruins' and *drubaža* 'small pieces, fragments';
Latv. *knāpt* and *knābt* 'to peck',
Lith. *slãstai* 'trap', Latv. *slazds*,
Latv. *sniekt* 'to give' and *sniegt*
Lith. *klusnùs* 'obedient' and *glusnùs* 'id.'
Lith. *kriaũnos* 'handle of a knife' and *griaũnos*
Lith. *kémbė* 'clothes peg' and *gémbė*
Latv. *pires* 'sheep's dung' and *bires*
Lith. *trenėti* 'rot, moulder', Latv. *trenēt, drenēt*
Lith. *kaũkaras* 'hillock, high ground', *gaũgaras* 'id.' cf. Croat. čuka, čukara
Latv. *kapana* 'haystack' and *gabana*

[15] Of course, in Balto-Slavic, the vowel *a is not diagnostic, since PIE *o and *a > BSl. *a.

075_102_Filologija_60_Matasovic_1.indd 95 19.01.2014 22:09:40

In some cases, words showing this alternation may be Uralic loanwords, or they may reflect the pronunciation of originally Baltic words by speakers of Uralic, who underwent language shift. However, it is certainly not the case that all of these words must be loanwords from non-IE languages. Moreover, the alternation of voiced and voiceless consonants is sometimes also found in words with perfectly convincing PIE etymologies:

PSl. *drabъ 'rag, cloth' (ULus. *draby* 'clothes', Ukr. *drab* 'poor man', Slk. *drabina* 'side panel on a cart'), Lith. *dróbė* 'cloth', Trautmann 61, ESSJa V:100—101. This is plausibly connected with PSl. *drapati 'tear, scratch' (Russ. *drápat'*, Cz. *drápati*, Croat. *drápati*, ESSJa V:101—102) < PIE *drep- (Gr. *drépō* 'pluck', perhaps also Ved. *drāpí-* 'mantle').

PSl. *dupa, *duplja 'hole, hollow' (Russ. dial. *dúpa*, Cz. arch. *doupa* 'hollow', Slov. *dúpa* 'hole, burrow', Croat. *dúplja* 'hollow', Pol. *dziupla* 'hollow') vs. Lith. *daubà* 'ravine, hole', Goth. *diups* 'deep', OE *dēop* 'deep' < PIE *d^{h}ewb-, ESSJa V:157—158.[16]

PSl. *dьrzъ 'bold' (OCS *drъzъ*, Russ. dial. *dérzyj*, Cz. *drzý*) vs. OPr. *dirsos* 'good', Lith. *drąsùs* 'courageous', Gr. *thrasýs* 'bold', ESSJa V:208f. Slavic *-z-* is unexplained, but it must be old, since we would otherwise expect the change of *s > *x by RUKI-rule.

PSl. *lupiti 'to peel', *lubъ 'peel', Lith. *laupýti* 'break', *lùpti* 'peel', Latv. *laupît* 'peel', Lat. *liber* 'bark of a tree, book', Alb. *labë* 'rind', Lith. *lúobas* 'bast', Latv. *luobas* 'peel', OPr. *lubbo* 'bast, plank', Lith. lubà 'plank'. LIV reconstructs PIE *lewp- on the basis of BSl. alone and does not discuss the relationship to *lewb^{h}-. The forms with *-p- could have been generalized from the sigmatic aorist (*lewb^{h}-s- > *lewp-s-), but there is no evidence that this verbal root ever formed a sigmatic aorist.

PSl. *trątъ 'drone', Lith. *trãnas*, Latv. *trans, tranis*; certainly related to Germ. *Drohne*, Gr. *thrõnaks,* but details of development are unclear (Smoczyński 682). Holzer (1989) interprets this word as "Tememathic" (see above), but the voiceless initial stop is the only argument for this hypothesis. In light of the overall uncertainty of the "Tememathic theory", it is better to simply acknowledge that we do not know the source and the original form of this word.

PSl. *tъp- / *dъp- 'to stamp': Pol. *deptać,* Cz. *deptati* vs. Russ. *toptát'*, Croat. *táptati, topot*; the original form had *t-, cf. Gr. *týptō* 'beat', Skr. *tupáti* 'id.', Latv. *staupe* 'trampling (of horses)', Vasmer III:122.

[16] Germanic *-p- might be explicable by Kluge's law, as arising from *-b- in a nasal stem.

075_102_Filologija_60_Matasovic_1.indd 96 19.01.2014 22:09:40

The causes of alternations of voiced and voiceless stops can often be accounted for within Indo-European. For example, the root-final voiceless stops in verbal roots can be the result of de-voicing before *-s- in the sigmatic aorist, and the root-final voiceless stop in a nominal form can be the result of generalization of the de-voiced stop which was regular before the nominative singular ending *-s in a root noun (Matasović 2011). Thus, it is methodically objectionable to derive all words in which Slavic and Baltic have unexpected articulation of stops from a non-IE substratum.

Here is, finally, a tentative list of Balto-Slavic words with no plausible PIE etymology, which could have been borrowed from some substratum:

Probable: *(a)rayHšu- 'nut', *bauKura- 'hillock', *e/amela- 'mistletoe', *graSa- 'terrible', , *Krawša- 'pear', *ledu- 'ice', *me/arG-ska- 'net', *śama- 'sheat-fish', *Traupa- 'log', *tulŚi- 'thick, fat'.

Possible: *PurHna- 'snout', *īni- 'hoar-frost', *Klawša- 'deaf, dumb', *seylā 'force', *tranTa- 'tree fungus'.[17]

Most of these words belong to semantic fields that are easily prone to borrowing. However, they do not share any obvious features that would help us attribute them to a single substratum. The irregular vocalic alternations and the alternations between voiced and voiceless stops are not specific enough, and we find similar alternations in the words belonging to the "Northwest European" layer of vocabulary in Celtic, Germanic, and Italic. Thus, we can conclude that there is no reason to assume that Balto-Slavic borrowed words from a single substratum.

Discussion and conclusions

The examination of the list of words belonging to the "European" vocabulary in Balto-Slavic presented in the last chapter allows us to make the following observations:

Firstly, Baltic and Slavic share this European vocabulary much more often with the western and northern European languages (Germanic, Italic and Celtic) than with the southern ones (Greek and Albanian), see Table I. This observation is consistent with the hypothesis, common among archaeologists, that Baltic and Slavic, together with Germanic, and possibly also Celtic and Italic, arose on the territory of the Corded Ware Horizon of the late 4th and the 3rd millennium BC (Mallory 1989, Anthony 2007:344—370). The "Northwest European" vocabulary was borrowed from substratum language(s) in the area occupied by the Corded Ware Horizon, which

[17] Capital letters indicate alternation between voiced and voiceless consonants, i. e. *S can be both *s and *z, *T can be both *d and *t, etc.

075_102_Filologija_60_Matasovic_1.indd 97 19.01.2014 22:09:40

was Indo-European, according to the communis opinio among archaeologists.

Secondly, words attested only in European branches of IE, but lacking in Anatolian, Tocharian, Indo-Iranian and (somewhat less commonly) in Armenian, are not often shared by both Baltic and Slavic groups of languages (see Table I). This can be explained in two ways. We might assume that substratum words from the Northwestern European substratums entered the ancestors of modern Baltic and Slavic languages after the Balto-Slavic period, when Baltic languages and Proto-Slavic were parts of a large dialect continuum stretching over much of Central and Eastern Europe. Only the extreme parts of that continuum were preserved until the present: the Eastern Baltic languages, and Proto-Slavic, which is relatively shallow, since it was spoken in the 5th century AD. The subsequent expansion of the Slavs covered much of the earlier dialect continuum, erasing many idioms previously spoken between the Proto-Slavic and the Eastern Baltic areas. During the time of the borrowing of non-IE loanwords, dialects belonging to different parts of this dialect continuum borrowed words from rather different substratum or adstratum languages. This would mean positing Proto-Balto-Slavic at a very early period, presumably before the Corded Ware Horizon in the third millennium BC. Considering how close the Balto-Slavic languages are from the dialectal point of view, I am inclined to believe the other possible explanation: that loanwords belonged to semantic fields in which rates of lexical replacement are very high, so that the original non-IE loanwords usually survived only in parts of the original Balto-Slavic area.

Thirdly, the number of words that may be of substratum origin, and that are preserved only in Balto-Slavic, is very limited (perhaps as few as 14, but probably not more than 20). It is significantly smaller than the number of words of substratum origin that can be attributed to Proto-Celtic, or to Insular Celtic (see EDPC), and it is also much smaller than the number of substratum words in Greek, for example.[18] This is probably due to the fact that, during the Balto-Slavic period, speakers of that proto-language were surrounded by speakers of other, more peripheral Indo-European dialects (especially Germanic and Celtic) that were exposed to more intensive contacts with speakers of non-IE languages. Consequently, during the period when Balto-Slavic separated from the other NW European dialects as an individual idiom, borrowing from non-IE substrata was minimal.

[18] Apparently as much as 10 % of Greek words in Beekes' recent etymological dictionary (2011) are of substratum origin.

075_102_Filologija_60_Matasovic_1.indd 98 19.01.2014 22:09:41

Words shared with NW European	*Words shared with SE European*	*Words with cognates in most European branches of IE*
26 (12 attested in both Baltic and Slavic)	4 (2 attested in both Baltic and Slavic)	10 (6 attested in both Baltic and Slavic)

Table I: The distribution of possible Substratum words in Balto-Slavic

References

Andersen, Henning. 2003. Slavic and the Indo-European Migrations. *Language Contacts in Prehistory. Studies in Stratigraphy* (Current Issues in Linguistic Theory 239) ed. by Henning Andersen, Amsterdam—Philadelphia : John Benjamins. 45—76.

Anreiter, Peter. 2001. *Die vorrömischen Namen Pannoniens.* Budapest : Archaeolingua.

Anthony, David. 2007. *The Horse, the Wheel, and Language.* Baltimore : Johns Hopkins University Press.

Beekes, Robert S. P. 1996. Ancient European Loanwords. *Historische Sprachforschung* 109, 215—236.

Boutkan, Dirk. 1998. On the form of North European substratum words in Germanic. *Historische Sprachforschung* 111, 102—133.

Boutkan, Dirk. 2003 On Gothic *MAGAÞ* ~ Old Frisian *MEGITH* and the form of some North European substratum words in Germanic. *Amsterdamer Beiträge zur älteren Germanistik* 58, 11—27.

Brozović, Dalibor. 1992. Review of Holzer 1989. *Wiener Slavistisches Jahrbuch* 35, 223—225.

Derksen, Rick. 2000. Old Icelandic *jarpi* 'hazel-grouse', *rjúpa* 'ptarmigan' and their Germanic and Balto-Slavic cognates. In: D. Boutkan and A. Quak (eds.) *Language contact. Substratum, superstratum, adstratum in Germanic Languages,* Amsterdam : Rodopi. 97—105.

Endzelīn, Janis. 1971. *Jānis Endzelīns' Comparative Phonology and Morphology of the Baltic Languages.* The Hague : Mouton.

Gołąb, Zbigniew. 1972. Kentum Elements in Slavic. *Lingua Posnaniensis* 16, 53—81.

Gołąb, Zbigniew. 1990. *The Origins of the Slavs.* Columbus : Slavica.

Holzer, Georg. 1989. *Entlehnungen aus einer bisher unbekannten indogermanischen Sprache im Urslavischen und Urbaltischen.* Vienna : Verlag der Österreichischen Akademie der Wissenschaften.

Kallio, Petri. 2005. A Uralic Substratum in Balto-Slavic Revisited. In:

075_102_Filologija_60_Matasovic_1.indd 99 19.01.2014 22:09:41

Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft. Ed. by Gerhard Meiser & Olav Hackstein, Wiesbaden : Reichert. 275—283.

Katičić, Radoslav. 1976. *Ancient Languages of the Balkans.* The Hague : Mouton.

Kortlandt, Frederik. 2003. An Indo-European substratum in Slavic. In: A. Bammesberger (ed.) *Languages in Prehistoric Europe.* Heidelberg : Winter. 253—260.

Kuiper, Franciscus B. 1995. Gothic *bagms* and Old Icelandic *ylgr*, *NOWELE* 25. 63—88.

Mallory, James. 1989. *In Search of the Indo-Europeans.* London : Thames and Hudson.

Matasović, Ranko. 2005. The Centum Elements in Balto-Slavic. In: *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft.* Ed. by Gerhard Meiser & Olav Hackstein. Wiesbaden : Reichert. 363—374.

Matasović, Ranko. 2007. Notulae etymologicae, *Folia onomastica Croatica* 15, 159—164.

Matasović, Ranko. 2011. The etymology of Latin *focus* and the devoicing of root-final stops before *s in PIE. *Historische Sprachforschung* 123, 212—216.

Matasović, Ranko. 2012. Areal Typology of Proto-Indo-European: The Case for Caucasian Connections. *Transactions of the Philological Society* 110, 1—28.

Moszyński, Leszek. 1992. Review of Holzer 1989. *Rocznik Slawistyczny* 48/1, 88—95.

Oettinger, Norbert. 2003. Neuerungen in Lexikon und Wortbildung des Nordwest-Indogermanischen. In: A. Bammesberger (ed.) *Languages in Prehistoric Europe*, Heidelberg : Winter. 183—194.

Schrijver, Peter. 1997. Animal, Vegetable and Mineral: some Western European Substratum Words. In: *Sound Law and Analogy, Papers in Honor of Robert S. P. Beekes on the Occasion of his 60th Birthday.* Ed. by A. Lubotsky, Amsterdam and Atlanta : Rodopi. 293—316.

Schrijver, Peter 1999. On henbane and early European narcotics. *Zeitschrift für celtische Philologie* 51, 17—45.

Stang, Christian 1972. *Lexikalische Sonderübereinstimmungen zwischen dem Slavischen, Baltischen und Germanischen.* Oslo-Bergen-Tromsø : Universitetsforlaget.

075_102_Filologija_60_Matasovic_1.indd 100 19.01.2014 22:09:41

Lexica

Beekes = R. S. P. Beekes. 2011. *Etymological Dictionary of Greek.* Leiden : Brill.

Bezlaj = F. Bezlaj. 1976—2007. *Etimološki slovar slovenskega jezika.* Ljubljana.

Chambers = R. K. Barnhart (ed.). 1988. *Chambers Dictionary of Etymology.* Chambers.

de Vaan = M. de Vaan. 2008. *Etymological Dictionary of Latin and the other Italic Languages.* Leiden : Brill.

Derksen = R. Derksen. 2008. *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden : Brill.

EDPC = R. Matasović. 2009. *Etymological Dictionary of Proto-Celtic.* Leiden : Brill.

DELG = P. Chantraine. 1967. *Dictionnaire étymologique de la langue grecque.* Paris.

EIEC = Mallory, J. P. & Adams, D. Q. *Encyclopedia of Indo-European Culture*, London-Chicago : Fitzroy Dearborn 1997.

EWA = M. Mayhrofer. 1992—2001. *Etymologisches Wörterbuch des Altindoarischen.* Heidelberg.

Gluhak = A. Gluhak. 1991. *Hrvatski etimološki rječnik.* Zagreb : August Cesarec.

IEW = J. Pokorny. 1959. *Indogermanisches etymologisches Wörterbuch.* Bern : Francke.

LIV = H. Rix *et alii.* 1998. *Lexikon der indogermanischen Verben.* Wiesbaden : Harrassowitz.

NIL = Wodtko, D. S., Irslinger, B. & Schneider, C. 2008. *Nomina im indogermanischen Lexikon.* Heidelberg : Winter.

Orel = V. Orel. 2003. *A handbook of Germanic etymology.* Leiden : Brill.

Skok = P. Skok. 1985. *Etimologijski rječnik hrvatskoga ili srpskoga jezika.* Zagreb.

Smoczyński = W. Smoczyński. 2007. *Słownik etymologiczny języka litewskiego.* Vilnius : Uniwersytet wileński.

Snoj = M. Snoj. 2003. *Slovenski etimološki slovar*. Ljubljana : Modrijan.

SP = *Słownik prasłowiański*, pod redakcją Franciszka Sławskiego. 1974—. Wrocław etc. : Polska Akademia Nauk.

Trautmann = R. Trautmann. 1921. *Balto-slavisches Wörterbuch.* Göttingen .

ESSJa = V. N. Trubačev (ed.). 1974—. Ètimologičeskij slovarʻ slavjanskix jazykov. Moscow.

075_102_Filologija_60_Matasovic_1.indd 101 19.01.2014 22:09:41

Supstratne riječi u baltoslavenskim jezicima

Sažetak

U ovom su članku prikupljene i analizirane riječi posvjedočene u baltoslavenskim jezicima bez jasne indoeuropske etimologije, koje po jasno definiranim kriterijima mogu biti posuđenice iz nekoga supstratnog jezika (pripadaju određenim semantičkim poljima, pokazuju neobične tvorbene osobitosti, itd.). Pokazuje se da većinu takvih riječi baltoslavenski jezici dijele s drugim indoeuropskim jezicima sjeverne i zapadne Europe (osobito s germanskim, u manjoj mjeri s keltskim i italskim), dok su znatno malobrojnije riječi supstratnog podrijetla koje usporednice imaju u jezicima južne Europe (grčki i albanski). Raspravlja se i o pretpostavci Georga Holzera o postojanju niza riječi koje su u baltoslavenski posuđene iz iščezloga indoeuropskog jezika koji on naziva »tememathskim«, no pokazuje se da je vrlo malo tih riječi posvjedočeno i u baltijskim i u slavenskim jezicima, te da one obično imaju uvjerljive indoeuropske etimologije. Zaključuje se da je u baltoslavenskome vrlo malo riječi koje su posuđene iz nekoga supstratnog jezika iz kojega istovremeno riječi nisu posuđivane i u drugim granama indoeuropskih jezika.

Ključne riječi: etimologija, supstrat, pretpovijesni jezični dodiri, baltoslavenski, indoeuropski

Key words: etymology, substratum, prehistoric language contacts, Balto-Slavic, Proto-Indo-European

075_102_Filologija_60_Matasovic_1.indd 102 19.01.2014 22:09:41